

ARI Contractor Report 97-25

**An Integrated Database of Unit
Training Performance: Description
an Lessons Learned**

**Bruce C. Leibrecht
BDM Federal, Inc.**

This report is published to meet legal and contractual requirements and may not
meet ARI's scientific or professional standards for publication.

August 1997

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited

DTIC QUALITY INSPECTED 3

19970825 151

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director



NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

DTIC QUALITY INSPECTED 3

FOREWORD

As part of its training research and development mission, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) conducts research regarding computer-based training management tools. This research aims to enhance the capabilities of the Army's combat units to accomplish complex training requirements effectively and efficiently. The ability to integrate and analyze training performance data from the diverse unit training arenas offers a powerful asset for planning and managing unit training programs. A comprehensive pool of data would also support the planning and allocation of training resources and facilitate the evaluation of the training value provided by complex training systems.

The ARI's Simulator Systems Research Unit has developed a prototype relational database for integrating unit training performance data. Funding for the project was provided by the Defense Manpower Data Center (DMDC). Their goal was to establish a comprehensive pool of data to support the planning and programming of training resources, including cost-benefit analyses of alternative training technologies. The DMDC's interest included the development of state-of-the-art methods for assessing the effectiveness of complex training systems. This report describes the structural and functional features of the database, including the analytical capabilities, and discusses the lessons learned during development. Recommendations for future research are also presented.

The prototype database establishes a foundation for expanding the data collection, processing, and analysis capabilities to support the Army training community. The potential to establish a comprehensive pool of training performance data promises to make valuable information available to unit trainers, training managers, training developers, policy makers, and researchers.

EDGAR M. JOHNSON
Director

ZITA M. SIMUTIS
Director, Personnel and Training
Research Division

**AN INTEGRATED DATABASE OF UNIT TRAINING PERFORMANCE:
DESCRIPTION AND LESSONS LEARNED**

	Page
INTRODUCTION	1
Background	1
Technical Objectives	3
INTEGRATED DATABASE DESIGN	5
Database Architecture	6
Data Analysis Capabilities	8
METHODS	10
Database Development	10
Data Collection	12
Data Processing	12
LESSONS LEARNED	14
Database Design	14
Database Development	18
Data Collection and Processing	20
Application Environment	23
Summary of Lessons Learned	24
Recommendations for Future Development	26
CONCLUSIONS	28
REFERENCES	29
ABBREVIATIONS	31
APPENDIX A. Measures of Performance	A-1

LIST OF TABLES

Table 1.	Integrated Database Design Principles	6
Table 2.	Kinds of Performance Data Defining the Start-up Integrated Database	7
Table 3.	Major Components of the Integrated Database	10
Table 4.	Summary of Lessons Learned	25

AN INTEGRATED DATABASE OF UNIT TRAINING PERFORMANCE: DESCRIPTION AND LESSONS LEARNED¹

INTRODUCTION

The U.S. Army's modern training environment is rapidly becoming more complex, particularly in terms of new training tools and changing mission requirements. Emergent training aids, devices, simulators, and simulations (TADSS) include distributed simulation networks based on virtual simulation technologies, such as Simulation Networking (SIMNET; see Alluisi, 1991). Other advanced TADSS include the Unit Conduct of Fire Trainer (UCOFT) and the Close Combat Tactical Trainer (CCTT), an application of Distributed Interactive Simulation (DIS) technology. Unit commanders face bewildering challenges in deciding which training tools will best meet their combat readiness requirements.

As part of its training research and development mission, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is conducting multi-faceted research to support Army training needs. A portion of this research aims to provide computer-based training management tools in the hands of unit trainers, training developers, and training managers. In pursuing this research, ARI's Simulator Systems Research Unit conducted a project designed as the initial step in creating a comprehensive training performance database. Focusing on training assessment among U.S. Army units in Germany, this project targeted database requirements in the SIMNET and field training environments. The central goal was to capture and archive measures of unit performance, mainly quantitative measures from automated training assessment systems. This report describes the prototype database system, including data capture mechanisms, and the lessons learned that may be of value to researchers in the future. It also suggests steps for expanding the scope of the database.

Background

Senior leaders and planners in the Department of Defense (DoD) and Department of the Army face difficult challenges in planning and resourcing training for combat readiness. Ensuring adequate training resources demands information on the relative effectiveness and costs of various options for training maneuver units and their staff elements. Today's training environment is extremely complex (U.S. Army Training and Doctrine Command, 1994) and includes large-scale interactive simulations such as SIMNET and the CCTT. Training packages harnessing SIMNET capabilities have been developed under ARI's Simulation-Based Multiechelon Training Program for Armor Units (SIMUTA; see Hoffman, Graves, Koger, Flynn, & Sever, 1995), and plans are

¹ Special thanks go to Dr. John Boldovici, who provided critical guidance and input to the planning and development of the integrated database. The author gratefully acknowledges the contributions of the following individuals: Jack Briscoe, David Butterfield, Joseph Cassidy, Kevin Clark, Timothy Clifton, Glenn Daens, Debrah Dornath, Stephen Goldberg, MAJ Kent Lasneske, Thomas J. Lewman II, Larry Meliza, Ray Miller, Halim Ozkaptan, COL James Rowan, Bruce Sterling, Bill Walsh, William West, Robert K. White, and Beverly Winsch.

underway to convert the packages for use in the CCTT environment. Still other training packages utilizing constructive simulations as well as SIMNET are being developed in ARI's Combined Arms Operations at the Brigade Level, Realistically Achieved through Simulation (COBRAS) Program (The COBRAS Team, 1995). The complexity of modern training systems makes it difficult to evaluate the training effectiveness of an individual training system in isolation (Boldovici & Bessemer, 1994). Because of the complexity, military trainers will require multiple training cycles to learn how to achieve the maximum potential of the entire system. Consequently, gauging the proficiency-enhancing contributions of the various TADSS and field training requires evaluating TADSS performance as part of the whole training environment.

Learning curve theory (Hancock & Bayha, 1982) holds that as the number of training repetitions increases and feedback is provided, the proficiency level of the trainees improves. Accuracy scores increase, performance speed accelerates, etc. The learning curve consists of two components, one related to the trainees as they perform the functions of interest, and the other related to the experience of the trainer(s) in administering the training process. According to the theory, each doubling of the number of training repetitions results in performance improvement by some fixed percentage. The rate of overall improvement is a function of the rates at which the individual trainees increase their proficiency levels. In quantifying the learning curve, performance data must be collected at several points during the course of training with new or conventional training systems (see Hiller, McFann, & Lehowicz, 1994).

To support the design, development, fielding, and continuing improvement of training systems, new methods are needed for evaluating their training effectiveness (Boldovici & Bessemer, 1994). Developing new systems requires comprehensive information about alternative training media mixes, including virtual simulation. The aim should be to relate the measured effectiveness of each media mix to the expected performance in a combat-like environment. Current training effectiveness evaluation methodologies typically use single-stage experimental designs. In this approach, the measured performance of the subject group represents a single point on the confounded learning curves of trainees and trainers as they adapt to the new system. Thus, single-stage methodologies yield inadequate data for judging the true potential of complex training systems (Boldovici & Bessemer, 1994).

In the mid-1980's the Army began developing training management and assessment systems capitalizing on micro-computer technology. The Integrated Training Management System (ITMS; see Madden, 1989) was developed to assist unit trainers in planning their training activities, documenting resource utilization, and tracking unit performance. The ITMS evolved into the Standard Army Training System (SATS; see U.S. Department of the Army, 1989), now the Army's principal training management system. On the performance assessment side, the National Training Center (NTC) implemented a performance database incorporating input from instrumented ranges. Similar systems have been installed at the other Combat Training Centers (CTCs)--the Combat Maneuver Training Center (CMTC) and the Joint Readiness Training Center (JRTC). In parallel, ARI established a CTC Archive to capture and store unit performance data collected during training at the CTCs. Data from the CTC Archive have been used, for example, to statistically profile performance of units completing NTC rotations. For the SIMNET training environment, ARI developed the personal computer-based Unit Performance Assessment System (UPAS) for use by

trainers and researchers (Meliza, Tan, White, Gross, & McMeel, 1992, 1994). The value of UPAS workstations in supporting after action reviews (AARs) has been documented by Meliza, Bessemer, Burnside, and Shlechter (1992), and the workstations have been incorporated in the SIMUTA training exercises. In a related effort, ARI is developing the Automated Training Assessment and Feedback System (ATAFS), a Unix-based training analysis tool for the DIS environment. And in the Electronic Collection Instrument (ECI) project, ARI has developed electronic clipboards to enhance the capability of trainers and observers in any arena to capture observations of task performance during execution of training exercises. The growing family of data collection systems sets the stage for systematically gathering training performance data to support analysis and research needs of trainers, training managers, training developers, and policy makers.

An important link in the Army's training assessment and management capabilities is the emerging Standard Army After Action Review System (STAARS). Currently under development, this system is being designed to provide standard tools for trainers to use in evaluating the effectiveness of their training exercises (U.S. Army Training and Doctrine Command, 1994). The concept calls for packages of AAR products standardized for each type of unit by echelon, regardless of the training environment. In the future, STAARS will gather information from all operational AAR systems. Significantly, the system is intended to collect standardized information from virtual, constructive, and live training exercises. The STAARS will interface with the developmental Army Training Digital Library (ATDL) to make information available to training, research, and development personnel.

The need for a comprehensive, integrated training effectiveness database is becoming more and more apparent. Trainers conducting SIMNET and CCTT training exercises need quantitative assessment tools for quick-response support of AARs and take home packages. The intelligent use of comparable performance data obtained in home station, training area, simulation center, and CTC training exercises is essential to units' planning and execution of successful training programs. Training managers need tools to help realistically select alternative training methods, based on expected quality of training and relative cost. Training developers and policy makers need accurate data to improve training systems and set realistic policy. These various needs point to a requirement for a central forum for integrating diverse training performance data. To ensure responsiveness to a broad spectrum of questions, the means to facilitate extraction of data for decision-making and research purposes is essential. Finally, the ability to input data from innovative platforms such as UPAS and electronic clipboards is important to support a comprehensive database.

Technical Objectives

The research described in this report was designed to meet four specific technical objectives:

1. To design and develop a prototype, multi-purpose, relational database for efficient storage, retrieval, and analysis of data on the kinds and amounts of training received by Army units. The database must permit the easy manipulation of unit training performance data from home station, training area, SIMNET, and CTC training exercises.

2. To populate the database with sample performance data from the U.S. Army, Europe (USAREUR) training domain.

3. To establish the analytical foundation for answering questions regarding training development, training policy, and research issues.

4. To establish an operational springboard for expanding the scope of the database, including the data collection network.

INTEGRATED DATABASE DESIGN

The prototype integrated database is designed as a central collecting point for multi-source performance data, as represented in Figure 1. The database comprises a relational system for integrating training performance data in the form of measures of performance (MOPs) or related data elements. The data originate from home station, training area, SIMNET, and CTC arenas, with the USAREUR domain serving as the starting point. The database is designed to answer questions posed by unit leaders, trainers, resource planners, training developers, and researchers. The personal computer (PC) based system encompasses data entry functions, organization and storage of data in a family of tables enabling easy manipulation of data, continuous monitoring of quality of the data, and retrieval and extraction of data for input to complex statistical analyses. The database accommodates currently available MOPs (Appendix A), while providing the capability to expand data inputs. Table 1 outlines the principles followed in designing the database.

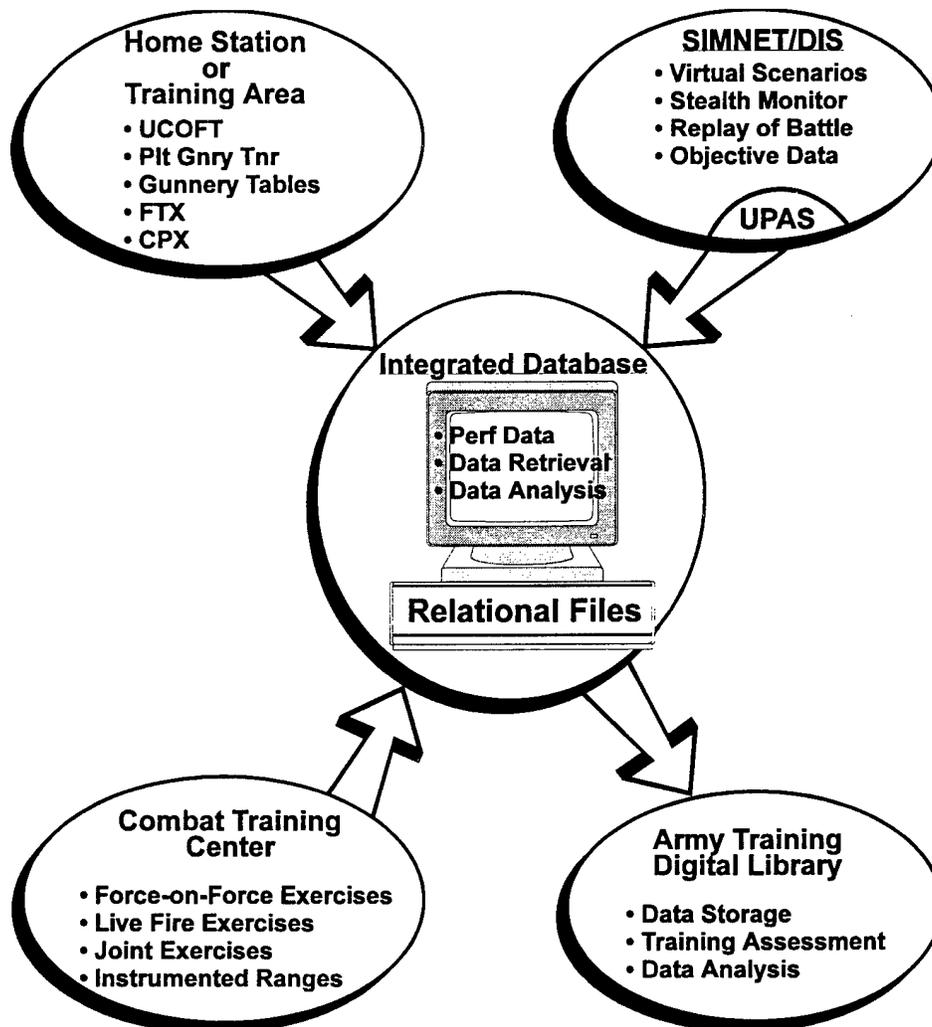


Figure 1. Role of the integrated database in Army training.

Table 1 - Integrated Database Design Principles

- Establish relational database on PC platform using commercial off-the-shelf software
 - Include available MOPs from home station, training area, SIMNET, and CMTC
 - Structure to enable rapid manipulation and extraction of data
 - Organize to support robust statistical analyses
 - Provide hooks for expanding data elements and sources
 - Establish central service point for data processing and customer service
 - Facilitate continuous monitoring of data quality
-

Database Architecture

The selection of MOPs for the integrated database began with a review of MOPs used by USAREUR units and CMTC staff, as well as MOPs obtainable from UPAS. These measures were analyzed against the following criteria: relevance to the technical objectives, utility to one or more user groups, practicality of accurate and consistent collection, and apparent validity. Measures meeting the criteria were incorporated in a list of MOPs to be supported by the integrated database. Table 2 lists the kinds of data included in the database; and a full listing of MOPs appears in Appendix A. It is important to note that the MOPs meeting the criteria dealt almost exclusively with weapons firing events and their outcomes. Unfortunately, MOPs related to command, control and communications were not available. Such process data are important to the interpretation of battle outcome data.

The integrated database is structured primarily to support those MOPs in Appendix A, based on readily available data elements. Briscoe (1996a) includes an inventory of the tables and their contents in his description of the database. Data tables are organized to conform to the general source of the data--home station, training area, and SIMNET--leading to families of tables related to data origin. The data tables receiving input via direct electronic transfer correspond to the structure of the source files. In the case of SIMNET data, a decision was made to incorporate unprocessed UPAS data files to conform with the structure familiar in the CTC environment and to provide maximum flexibility of supportable MOPs. The data tables receiving manually input data are organized, in part, to facilitate the keyboard entry process. An expandable table structure enables new data to be added to previously accumulated data, thus integrating data elements across units and across successive exercises. Data tables are not segregated by size and type of unit; rather, each data set carries codes to mark the unit's size, type, nature of mission, and other identifying information.

Table 2 - Kinds of Performance Data Defining the Start-up Integrated Database

Training Locus	Type of Training	Data Source
Home station	Tank UCOFT (crew)	7ATC Files
	Bradley UCOFT (crew)	7ATC Files
	TOW/DRAGON gunnery (crew)	7ATC Files
Local/major training area	Platoon Gunnery Trainer (Tank)	7ATC Database
	Platoon Gunnery Trainer (Bradley)	7ATC Database
	Tank Gunnery Table VIII	7ATC Database
	Tank Gunnery Table XII	7ATC Database
	Bradley Gunnery Table VIII	7ATC Database
	Bradley Gunnery Table XII	7ATC Database
	Aviation Crew Gunnery Table VII	7ATC Database
	Aviation Crew Gunnery Table VIII	7ATC Database
SIMNET Center	Platoon/Company/Battalion Mission	UPAS
CMTC	Battalion/Task Force Mission	CMTC Database

The database is fully relational--any field in any table may be related to one or more fields in any other table within the database. This feature enables data elements from different sources to be linked for analytical purposes. Thus, data from disparate tables can be combined in a common analytical framework where identity of specific crews or units is a critical factor.

Only the minimum number of tables required to accommodate the available data have been established initially. Because resource and other constraints precluded collection of CMTC data, tables for those data were not constructed in the start-up database. New tables can be added as additional sources of data are identified. A master inventory of data elements is maintained by the operating staff, indicating where a given element can be located within the set of database tables.

The functional features of the integrated database include input, output, data conditioning, and query or report functions (Briscoe, 1996a). The following features are available:

- Collection of training performance data in electronic and paper forms
- Data conversion and entry functions, both electronic (tape, diskette) and manual
- Reformatting functions to standardize the form of the data elements
- Mass storage (internal and external) of data organized in source-linked tables
- Display of data on screen

- Preparation of queries and reports, including manipulation of data
- Output of tailored data files on floppy diskette or tape
- Printout of data, queries, and reports

Data Analysis Capabilities

The prototype database supports analysis of training performance data, although its utility is constrained by the lack of data beyond firing events and battle casualties. Direct analytical functions include queries and reports yielding descriptive information. In addition, custom data files can be provided to clients interested in conducting their own statistical analyses or modeling. Alternatively, a central analysis capability could be established to provide customers with full-service analytical and statistical support.

A query is a means of interrogating the database directly to return descriptive data related to a specific question of interest. Data elements must be designated by table and field name. Because of the relational nature of the database, data elements from multiple tables may be included in a query. As a simple example, an investigator could examine the ranks of Bradley commanders who obtained Table VIII scores exceeding 950. Queries can be used to combine data elements or otherwise compute desired MOPs. Query results can be output in the form of a printed table, a file on tape or diskette, or an intermediate table for further processing.

A report is a query-like tool for interrogating the database, but it offers expanded formatting capabilities available through the report writer features of Access®. A report uses the results of a query as its basic input, allowing the analyst to reorganize the data elements for greater clarity or ease of interpretation. Thus, reports afford the same analytical power as queries. Questions addressable by means of queries and reports tend to be simple and descriptive in nature, representing situations where a straightforward listing of data elements suffices. For example, an analyst might want to select those platoons with a maximum total score on Tank Table VIII and determine their speed and accuracy scores on Platoon Gunnery Training.

The integrated database's resident data manipulation and retrieval capabilities can generate output files containing data elements as specified by a customer. The data files can be structured for input to commercially available statistical analysis software, such as the Statistical Package for the Social Sciences (SPSS®). The output files can be written to floppy diskette or tape for transfer to the analysis system. Examples of supportable analytical approaches include *t*-tests, analysis of variance, regression analysis, discriminant analysis, nonparametric correlation or inferential models, and similar statistical treatments. Requests for such data files must be clearly and precisely stated in terms that can be translated into data parameters contained in the integrated database.

A question of interest at the outset of this project concerned the relation between performance at the CMTC and performance in pre-CMTC training exercises. In particular, we planned to identify predictors of success at the CMTC. The significance of the question lies in the practical value of knowing what training events units can emphasize to improve their readiness for CMTC exercises. The original plan was to examine MOPs from home station, training area, and SIMNET exercises to determine which measures are useful predictors of CMTC performance. Multiple linear

regression models were to be developed for analysis of cumulative platoon performance data, including derived variables. When CMTC data did not materialize, the analysis was deferred, pending further development of the database. However, the planned analysis illustrates the kind of complex, meaningful question addressable with the integrated database.

Intelligent analysis and interpretation of data from the integrated database require a clear understanding of the raw data and where they come from. Knowledge of the MOPs and the conditions under which they were collected is essential. Also important is an understanding of the structure of the tables in the database and the relations among them. Finally, robust analytical capabilities require later expansion of the MOPs to include data illuminating command, control, and communications processes.

METHODS

Translating the database's design features into an operational system involved several stages. The research team developed the functional software capabilities using a simple applications development approach built around a commercial database package. Collection of data for populating the database entailed establishing routine data gathering channels for the target training arenas. Finally, simple data processing procedures were adapted for screening, inputting, and checking the collected data.

Database Development

The integrated database is organized to assemble, condition, and process data, with associated input and output functions. Five components are involved: data collection, input, standardization, archiving, and analysis components. Table 3 summarizes these components.

Table 3 - Major Components of the Integrated Database

Component	Functions
Data Collection	Gathering of multi-source data under standard conditions Packaging of data and shipment to central collecting point
Data Input	Electronic conversion/transfer of tape and diskette files Manual keyboard entry of paper-based data
Data Standardization	Conversion of data to standard units of measure Consolidation or reformatting of data to standard specifications
Data Archiving	Repository of data in unified database Maintenance of data integrity throughout database life cycle
Data Analysis	Preparation of queries and reports (descriptive analysis) Output of custom files for statistical analysis

The front-end engine for the database--the data collection component--was developed by leveraging existing sources of training performance data and by adapting the UPAS workstation (Meliza, Tan, et al., 1992) for the SIMNET environment. For home station and training area data, the research team made arrangements with USAREUR's 7th Army Training Command (7ATC) to periodically obtain copies of selected data (see Table 2) maintained by the command's Training Analysis Division. File format specifications were agreed upon, and packaging and shipping procedures were established.

To enable collection of standard performance data in the SIMNET training arena, computer-based tools were established by adapting ARI's UPAS workstations. The basic functional capabilities of UPAS have been described by Meliza, Tan, et al. (1992, 1994). The workstations were enhanced by adapting new software developed under the SIMUTA project to make the workstations more user friendly. The SIMUTA software affected only the working user interface in order to make selected capabilities readily available for the expected user. Four workstations were installed in the SIMNET facility at the Grafenwoehr Training Area in Germany. This facility had the simulation resources to run platoon, company, and battalion training exercises, serving primarily two operational combat divisions.

In preparation for SIMNET data collection, the research team demonstrated UPAS capabilities and provided training for SIMNET site personnel. These sessions prepared the staff members to operate the workstations, to save data, and to assist unit personnel. In support of this training, a UPAS training package and a UPAS operating guide (Wunsch & Cassidy, 1996) were developed for the enhanced workstations to enable collection of high-quality data.

Existing data channels were targeted for collecting data from CMTC training exercises. Both quantitative performance data based on instrumented ranges and observer/controller observations were desired, along with supporting scenario materials. The Center for Army Lessons Learned (CALL) was considered a candidate to forward available data packages to the integrated database site. As mentioned earlier, resource and other constraints precluded CMTC data collection.

Excluding the data collection component, the remaining components of the integrated database were developed using primarily commercial off-the-shelf software--Microsoft Access® 2.0 on an IBM-compatible personal computer (Pentium® 90 MHz processor running under Windows for Workgroups 3.11®). The data input component supports both electronic and manual entry of data, as well as verification of data accuracy. In addition to the Access® software on a database workstation, the data input component includes a UPAS workstation for converting and screening UPAS data. The data standardization component harnesses Access® functions for consolidating or reformatting data to ensure consistent definitions and formats across the complete set of tables. The archival component consists of a number of data tables organized in a single relational Access® database. This component ensures that, as new data are entered, the integrity of previously entered data is protected. Finally, the data analysis component uses Access® query and report functions to prepare descriptive data summaries and to develop custom files to feed subsequent analyses using statistical analysis packages.

The key to developing the data processing components of the integrated database was the establishment of the table structure and data input schemes. Microsoft Access® software was used to create a set of data tables based on the form and characteristics of the input data received from each source. For simplicity of data entry and tracking, database tables were created to correspond directly with the tables received from the various sources. Tables were constructed in open-ended form to allow entry of data from sequential training exercises, given the continuing input of data. The research team then devised an import/implementation scheme for each source, using one of three types: (a) direct electronic import; (b) conversion (reformatting) of data followed by electronic import; and (c) import by means of keyboard entry. This step led to the establishment of the data

input component. Relations between tables were then examined to identify common fields and requirements for consolidated fields. This step led to the development of the data standardization component, based largely on measurement and formatting rules. Establishment of the data analysis component required primarily the verification of Access® functions for executing queries and reports, using sample data.

The database characteristics, including functional features and data table structure, were documented in a functional description (Briscoe, 1996a). In addition, user documentation (Briscoe, 1996b) was developed, explaining the functional features, input/output features, administrative procedures, maintenance and up-keep requirements, etc.

Data Collection

To initialize the database, preliminary data were collected in conjunction with scheduled training activities of USAREUR units, to establish the foundation for future expansion. The data collection period extended from April 1995 to November 1995. Data from home station, training area, and SIMNET arenas were obtained through separate pipelines, described in the following paragraphs.

Data collection in the SIMNET environment relied on Grafenwoehr site support personnel, who operated the UPAS workstations. At the end of each exercise, site personnel transferred UPAS data tables (unprocessed, in XDB format) to magnetic tape cartridge. Supporting materials such as operation orders (OPORDs) and event lists were to be collected as available. Accumulated tape cartridges and supporting materials (when available) were shipped periodically to the integrated database site, with a target of every two weeks. Throughout the data collection phase, the research team provided CONUS-based consultation and problem-solving services, including maintenance of a reference UPAS workstation at Fort Knox, where access to a SIMNET site was available.

Data files provided by 7ATC were written to floppy diskette in dBASE® format and shipped to the integrated database site every three months or so. All training exercises represented in the 7ATC database since the last shipment were to be forwarded. For data from home station training exercises (see Table 2), 7ATC personnel provided copies of paper printouts from their unit performance files. These materials were to be shipped on the same basis as the diskette files. The data obtained from 7ATC sources originated from exercises using standard training and measurement procedures.

Data Processing

Sample data gathered during the course of the project were entered into the integrated database, with data organized according to the established file structure. This included home station, training area, and SIMNET-UPAS data. The research team accomplished data entry and quality control, managed file configuration, and maintained database records.

Data from all sources were received at the integrated database site for processing. Data packages were inspected for missing data and apparent discrepancies, then readied for input to the

integrated database. Files received in dBASE® format (i.e., gunnery tables from the 7ATC database) were imported directly from diskette or tape into the Access® database, table by table. Files received from UPAS in XDB format were first converted to dBASE format using a UPAS workstation, then imported into the Access® database. Data in the form of paper records were tagged for entry by a keyboard operator, using input screens developed specially for the database.

Quality assurance procedures served to maintain high quality of the data in the integrated database. For UPAS data, preliminary review of an exercise in playback mode is often useful to identify potential data interpretation problems (e.g., three platoons conducting separate exercises but recorded as a unified company exercise). Data entered by direct electronic transfer can be inspected on-screen to verify that the transfer process worked properly. Keyboard entry of paper-based data requires verification of printouts against the original data received. For routine operations, spot checks of randomly selected data segments are planned for all kinds of data. All errors discovered are flagged for prompt correction.

LESSONS LEARNED²

Throughout the course of the project, the research team gleaned conclusions and lessons learned that may be of value in future efforts to develop or extend training performance databases. The lessons derive mainly from accumulated observations of members of the research team, but they also reflect input from data collection site staff and unit personnel. This section presents the major lessons learned, organized around the following topics: database design, database development, data collection and processing, and the application environment. A final subsection suggests some directions for future research.

The lessons learned should be considered in the context of the DoD's materiel acquisition paradigm. The development of a new training performance database may be governed by the procedures and strategies that apply to any military hardware/software system. Of particular significance is the Army's Manpower and Personnel Integration (MANPRINT) program, which defines steps for systematically addressing personnel, human factors, training, safety, and related dimensions of military systems. Many of the lessons discussed in this section pertain directly to MANPRINT requirements.

Database Design

Define the Target Audiences

An early step in the design process is the definition of key groups of people, or stakeholders, who will be associated with the database. Three specific audiences are typically defined: customers, operators, and maintainers.

1. Customers are those who use the products of the database, such as unit trainers, training developers, and researchers. They may use the products as is, or they may use the products for additional processing and analysis. Some customers may use exercise-specific products as data are collected, as is the case for trainers using immediate UPAS output to prepare for AARs. It is essential to define how the customers will use the database products. Customers are the key in determining what goes into the database.

2. Operators are the direct users of the database--those who operate the various components. Several groups of operators may be involved, depending on the number of data sources. Operators largely determine the design of the user interface(s).

3. Maintainers are responsible for upkeep of the database, including supply, maintenance, troubleshooting, and repair. The definition of maintainers is a key in determining logistical support requirements.

² Lessons learned in this report are discussed in generic terms, to enhance their generality and avoid attribution of responsibility outside the immediate research team.

The designation of the direct operators has special significance in the design process. The design of a database destined to be operated by a small, technically sophisticated group will be much different from that for a diffuse group of diverse personnel situated at dispersed locations. Discrete database components may target separate operator audiences, and this circumstance may drive disparate design criteria for different components. In the current project, the data collection elements were physically much different from the other components and located far away from the central database site. The definition of operator groups heavily influenced the planning of training capabilities and technical support mechanisms, as well as operational requirements. Characteristics of the operator audience help determine desired user interface characteristics, which can vary greatly for naive operators compared to personnel experienced in computer operations.

Meet the Needs of All Target Audiences

All three audiences--customers, operators, maintainers--are important in designing the capabilities and characteristics of the database's various components. The membership and needs of these audiences should be defined in detail. The more clearly these audiences are defined, including their respective roles in operating the database, the more nearly the operational database can meet their needs. The needs, preferences, and characteristics of each audience influence the content of the database as well as its functional capabilities. It is highly desirable for representatives of each audience to participate actively in the design process.

To ensure a demand for the products of the database, the expressed requirements and desires of potential customers should be taken into account. Preferably, customer representatives participate directly in the database design process. When that is not feasible, questionnaires, interviews, and similar instruments can gather information about customer needs and desires. Unit trainers are an important subgroup of customers. Unit trainers using structured training support packages (TSPs) available from programs such as SIMUTA need to collect MOPs that support AAR procedures built into the TSPs. The data collection components (e.g., UPAS workstations) of a training performance database should meet trainers' data collection requirements. In turn, the analytical capabilities of the database should link to the MOPs used for AAR purposes.

To ensure operational acceptability, particularly for the data collection components, the characteristics and preferences of operators should be addressed. Different groups of operators have different requirements, so each group should be addressed separately. Finally, the qualifications and needs of targeted maintenance personnel must be considered. As with operators, different groups of maintainers may be defined for different database components. In that case, each group should be defined and addressed individually.

The requirements definition process may address any or all of the following parameters: definition of customers, designation of target user populations, required MOPs, data input options, data manipulation and processing capabilities, data retrieval and extraction options, data storage and archiving needs, database output and display options, system flexibility and expandability, rules and means for customer access, user interface characteristics, backup and recovery requirements, compatibility with other platforms, security considerations, support constraints, and administrative and housekeeping requirements.

Understand Targeted Data

Beyond global statements of purpose and structure, detailed database design requires accurate descriptions of targeted data. Where existing data collection mechanisms are tapped, it is important to obtain complete data descriptions as early as possible to ensure a realistic design approach. Supporting information about the data collection environment is highly desirable. Where new data collection mechanisms are being created, early "paper" data descriptions may have to be supplemented with functional descriptions obtained during initial developmental trials. In the case of new data collection procedures, supporting information about the data collection environment will most likely have to wait for initial trials. Early coordination with cooperating agencies should be accomplished to ensure eventual availability of data.

Ensure Data Collection Complements Training

Not only should data collection tools be useful to users and unit trainers, they should not interfere with the unit's accomplishment of its training mission. They should fit comfortably in the operational training environment. Data collection components should be nonintrusive--they should not interrupt or delay the training exercises spawning the data. This is an important point in the database design process. The operational environment should be analyzed carefully to ascertain initialization concerns (e.g., who can provide unit and exercise identification data?), timing requirements (e.g., how quickly do trainers need end-of-exercise results?), physical constraints (e.g., how many soldiers need to see the display?), and security considerations (e.g., what safeguards are needed to prevent accidental failures?). Data collection tools that interfere with the training will almost surely sit unused at the training site.

The time required to implement the data collection procedures is important, as is the difficulty or complexity of the operations involved. This is not usually a problem where established procedures are harnessed, as long as additional steps are not demanding. However, where new data collection procedures are instituted, there is a potential to slow down the training exercises themselves. Especially when new data collection tools support a training exercise, the procedures must meet the time requirements of the training unit. For example, only ten minutes may be allowable for a component such as UPAS to generate MOPs for platoon and company AARs. If the MOPs are not ready in time, the immediate incentive for trainers and operators to use the component disappears. As a general rule, procedures should at least keep pace with the flow of the unit training exercises. In addition, the simpler the operating procedures, the better. If a new component is too slow or too complicated to keep pace, steps should be pursued to upgrade the operating speed and/or simplify the procedures.

Ensure Operator Incentives

Training performance data collection tools such as UPAS should provide tangible value to operators--those who are responsible for inputting the basic data or operating the devices so accurate data are recorded. This may well mean that tools must incorporate functions of immediate value to the operators and their customers--in most cases unit trainers--as well as data capture functions of interest to researchers, training developers, policy makers, etc. The importance of utility at the data

collection point should not be overlooked as goals related to integration and statistical analysis of data are pursued. In a word, the operators and trainers should be able to see an immediate payoff as well as a longer-term value to inputting accurate and reliable data.

Plan for Personnel Requirements

Based on the definition of operator and maintainer groups, planning and programming the personnel resources required to operate and support the database are essential. This is true for new data collection components, such as UPAS, and for data processing components of a central database. Even when common training and assessment procedures are followed, obtaining performance data from individual units can be difficult, considering the additional workload that may be imposed on unit personnel to collate and ship data packages. Careful assessment of the number of personnel, their qualifications, and manhours required per operational period must be accomplished. Some requirements may be serviceable by existing personnel, but others may mean establishing new positions. Positions can be identified or created in unit authorization documents of government organizations, or they can be programmed through contract mechanisms. If operators and maintainers are not slated for "ownership" by the developing organization, the program sponsor may well need to take the lead in arranging for personnel support. When a government agency provides support personnel, a memorandum of agreement between sponsoring and supporting agencies may be in order to spell out mutual responsibilities and resource arrangements. The personnel programming process should be timed to make operators and maintainers available at the start of database operations.

Plan for Logistical Support

An adjunct to the design process is planning for logistical support of the target database. Who will maintain the database and provide supplies? What replacement parts are to be stocked on-site? What troubleshooting is to be accomplished by operators? What is a reasonable technical support approach? How will minor and major repairs be accomplished? The logistical support plan will heavily influence the viability of the database. It provides the foundation for designating responsibilities, channels, mechanisms, and funding sources required to keep the database operating throughout its intended life cycle.

Ensure User Friendly Interface(s)

The type of user interface selected for database operations can influence strongly the user friendliness of the system. A text-based interface may be acceptable for users with a strong computer background. Such an interface often requires the operator to remember a basic set of operational commands, or to type in commands selected from a list of prompts. However, a graphical user interface--such as that used in Microsoft Windows®--assists the user with icons and pull-down menus for mouse selection, reducing the demands on the operator. The graphical interface is the type now used almost universally in desk top computers, so it enjoys familiarity among a broad base of potential users. Significantly, the graphical interface lessens the chance of operator errors introduced by typing command strings on the keyboard. A lower risk of operator

errors can impact positively the quality of data collected and/or entered. As a general rule, the graphical user interface is preferred in today's operating environment.

Plan for Multi-Phase Design Contingency

Relying on existing data sources, such as 7ATC's performance database, can facilitate the design of training performance databases and speed the collection of actual data. At the same time, such reliance limits the kinds and amounts of data collected and can exclude certain kinds of training exercises from the scope of the database. For example, home station and local training area exercises which utilize manual assessment techniques do not generate ready-to-gather data. (As the Army upgrades the SATS and moves towards universal usage of that system, availability of data in the field should improve substantially.)

Where new data collection mechanisms are being created, detailed design of the corresponding data processing components may well have to wait until the new collection procedures are in place or at least clearly specified. Thus, building a database around both existing and new data sources may mean that design of data tables and loading schemes must proceed in stages or in parallel with development activities. As selection and definition of data elements proceed, redundant and unnecessary data should be weeded out. It is important to emphasize the connectability of data from different sources throughout the design process. The upshot of a multistage design approach is two-fold: it complicates the design process, and it stretches out the time required. Project managers should take these impacts into account when planning resources, coordination requirements, and schedules.

Document the Design

An important product of the database design efforts is a document summarizing the training audiences, required capabilities, interface characteristics, targeted data, personnel and logistical support requirements, etc. Representing a functional description, the document can serve as a useful tool for coordination with various groups of stakeholders. It also provides a road map to guide development activities. As development activities progress, changes in the database design may become necessary. If substantive changes occur, it is wise to update the functional description to facilitate coordination and concurrence by the major stakeholders. This can be accomplished by issuing an addendum when only one or two components are modified, or by reissuing the entire document when many of the components are impacted.

Database Development

Development of the database should be driven by the outcome of the design process, as documented in the functional description or similar document.

Obtain Target Audience Input

As much as practical, the research team should obtain input and feedback from customers, operators, and maintainers as development of data collection and processing components proceeds.

Points of contact with target groups can serve as conduits for routine coordination and exchange of information. Routine communication will help keep the team up to date on changes occurring at the operating and customer sites. Documentation of emerging components can be circulated for review and comment. Visits to customer sites, data collection points, etc. can yield valuable insights, especially regarding the operating environment. Periodic in-process reviews can provide forums for informing stakeholders of progress and obtaining input on various issues encountered. Such steps represent quality review checks that can lead to enhanced customer and user satisfaction.

Sample the Targeted Data

Database development requires sample data from targeted sources. This acid test principle is an imperative, and it may necessitate developing data tables, conversion procedures, etc. in stages as different sample data packages become available. Along with the technical characteristics of the data, the operational nature of the data collection environment should be weighed. Of course, coordination with cooperating agencies typically must precede the delivery of data samples. As a general rule, detailed development and testing of data processing components must wait until sample data are in hand. When new data collection mechanisms are being installed, development of the data processing components may well be delayed pending the availability of data from all sources. On the other hand, an alternative is to begin development of components for existing data sources as soon as practicable, while waiting to develop the components for new data sources. A multistage approach can complicate the development process and extend the time required.

Verify the Feasibility of MOPs

As knowledge grows regarding the collectable data from both existing and new sources, the research team should routinely crosswalk the data elements with the list of targeted MOPs. This process helps determine if all desired MOPs are supportable. Samples of actual data may exhibit characteristics different from those advertised in the descriptive information typically obtained during the design phase. If the collectable data fail to support certain MOPs, the most likely option is to modify the list of targeted MOPs and coordinate the changes with the database stakeholders. It may be possible to redefine some of the MOPs in question, whereas others may have to be deleted. Another option may be to alter the data collection procedures to meet the specified MOP requirements. This option is especially difficult where established data collection procedures are involved.

Meet Operator Expectations

Selection of hardware platforms, software packages, user interface characteristics, output capabilities, and other implementation cornerstones should take into account the designated operators and their operational environments. This becomes a question of meeting the expectations of the operators and capitalizing on their habits. If the intended operators are most comfortable with personal computers running in a Windows® environment, the developers should strive to incorporate Windows® features. When an older system is being adapted as a component of the database, steps should be taken to upgrade the system so it is consistent with equipment and software in current use, to the extent affordable.

Test Components Thoroughly

As an integral part of the development process, the research team should conduct internal testing of all database components to ensure they work as designed. Interim testing is normally needed at the subcomponent level, followed by functional testing of complete components. The team should also perform verification tests of data collection components to ensure the collected data are accurate. This ground truth testing is critical to the credibility of the data collection tools. Such testing must be performed in a realistic operating environment, with network and terrain database characteristics matching those expected at the actual data collection sites. Problems should be documented so corrective steps can be taken. Calendar time for subcomponent, functional, and verification testing must be planned, and supporting facilities and personnel must be scheduled.

Plan Installation of Data Collection Components Early

Planning and coordinating the installation of new data collection components can be a challenge. The operating schedules at the data collection sites typically are quite constrained, and identifying a calendar window suitable for installation activities may be difficult. Substantial lead time may be required, and even then unexpected schedule changes may occur--both at the developing site and target installation sites. Close, continuous interaction between the research team and the data collection sites is imperative. A single point of contact for each group is a distinct advantage in this coordination process.

Have Developers Install Components

Wherever possible, new hardware and software should be installed by members of the research team who developed the new components. This is preferable to handing off the installation duties to personnel who did not participate in development. A critical part of the installation process is the on-site evaluation of hardware and software to verify that the component is working properly. This evaluation should begin with on-site functional testing and proceed to realistic operational testing with the site network fully loaded, preferably in a full-scale demonstration. Where extensive issues are of interest, more highly structured formative evaluation methods can be employed. On-site testing should include exercising the logistical support network as fully as feasible. If bug fixes and system adjustments require later delivery of a finalized component, the follow-up installation should include realistic evaluation. Upon completion of the evaluation, the research team should be convinced that the component is fully ready to begin routine data collection operations.

Data Collection and Processing

Ensure Motivated Data Collectors

Steps should be defined and implemented to ensure that personnel at the data collection site place high priority on collecting quality data. This helps establish data collection as a concentrated, focused effort yielding optimal inputs to the database. A memorandum of understanding between sponsoring and supporting agencies may be of value in establishing visibility and high level emphasis on the data collection effort. The issue of adequate staffing to support data collection

efforts, discussed earlier, becomes paramount here. Qualified operators and maintainers must be on hand, and they must be able to spend sufficient time to implement the specified data capture procedures. Routine, on-site supervision of ongoing data collection efforts is essential. The supervisor should be familiar with the data collection requirements and committed to collecting high quality data.

The success of a continuing database depends on how well operators and unit trainers accept the data collection efforts as valuable to them. Generation of high quality data requires proper use of the data collection components on a routine basis. Where new data collection components are introduced, deliberate steps are generally needed to promote their acceptance in the field. Arranging for an influential sponsor to proactively endorse the database effort can greatly facilitate acceptance at the working level. At least one supportive sponsor in each training domain is desirable. The sponsor can transmit messages to training units stressing the importance of the program and encouraging their support. A key point of contact can serve as a liaison between the sponsor and field units, providing guidance, information, and spokespersonship. Equally important is education of trainers and operators on how the database will benefit them. Information packages, orientation sessions, and train-the-trainer sessions are common means for educating key personnel. Where unit trainers do not operate the data collection component(s), the actual operators are influential in shaping the attitude of unit personnel. Operators who are convinced that the database adds value to the training exercises will facilitate acceptance of the system among unit trainers.

Priority of effort should extend to assembling, labelling, packaging, and shipping the data. Procedures for accomplishing these functions should be simple and error resistant, while maintaining accurate identification of the data. Clear instructions, to include the routine shipping schedule, should be prepared and placed in the hands of the responsible personnel. It may be possible to use Internet channels for delivery of data, such as E-mail or file transfer protocols. The on-site supervisor should ensure proper and timely execution of these functions.

Standardize the Data

Standardization of data inputs can be difficult or impossible to achieve, especially where training is conducted in the absence of standard exercises, as happens often in SIMNET training. Differences across training sites and units in terms of mission details, organization of friendly and enemy forces, quantification procedures, etc. can render data from similar exercises noncomparable-- a case of apples and oranges. Both training methods and performance assessment procedures must be standardized for data to be comparable across units and time. Training development initiatives such as the SIMUTA and COBRAS programs offer promise for standardizing collective training exercise procedures as well as the associated measurement methods. The SIMUTA exercises are already in use in CONUS, and an initial set has been exported to Germany.

When training exercises do not typically follow standard scenarios and procedures, information characterizing each exercise is important for interpretation of data. On-site personnel should gather supporting information from unit trainers, including type of mission, commander's intent, units involved, and enemy composition. The supporting information should be labelled clearly by date/event/unit and packaged to accompany the primary data. Without a doubt,

dependence on supporting information makes interpreting unit performance data much more difficult. This reinforces the importance of efforts to develop and field standard training packages for large scale simulations, such as the SIMUTA and COBRAS programs.

Provide Qualified Operators

Collection and processing of quality data require qualified and trained operators for the various database components. Realistic qualification criteria should be established and enforced in selecting operators. The next step is to train the operators in a fully operational environment. This will most likely necessitate the development of a "how to" training package, capable of being implemented by site personnel. Practical exercises should be included to ensure working knowledge of operations and afford the trainers an opportunity to assess the need for remediation. As data collection proceeds, refresher training may be needed, depending on the duration of the project and frequency of data collection cycles. When replacement personnel enter the picture, they must be thoroughly trained, too. Where resources permit, it is generally worthwhile to embed tutorials and "Help" routines in the operating components. This is true for all data collection and data processing components of the database.

Protect the Quality of the Data

Periodic visits to data collection sites by research team members represent a valuable quality assurance mechanism. Such visits can yield insight into operational problems, perhaps redefining technical issues. They can also create a two-way dialog to keep data collection personnel abreast of issues and help the research team identify new operational needs. The frequency of such visits will likely depend on the frequency of problems and the complexity of the operating environment. Planning and budgeting for site visits should be an integral part of project management.

Before entering data into the database, screening of all data is essential to determine if potential problems exist. Knowledgeable personnel should review both primary data and supporting information to ascertain if data are complete, properly aggregated, consistent with mission characteristics, etc. To support such screening, the research team should develop and disseminate criteria, working closely with data collection personnel and unit trainers. It is desirable to have the gross screening done at the collection site, by personnel most familiar with the data and the collection environment. Second-level screening can be performed by personnel at the database entry site. Such a two-stage screening process may be more effective than a single stage process. In addition to protecting the basic quality of the data, the screening process should be designed to eliminate redundant and uninterpretable data. Wherever possible, qualitative and quantitative tools should be made available to facilitate screening, such as UPAS playback procedures or preprocessing software.

The research team should routinely monitor the collection and processing of the data, with an eye on both quantity and quality of data. When problems are detected, it is important to provide feedback to the personnel involved so they can correct oversights and misunderstandings. In addition to feedback, supplemental instructions and/or training may be called for. Problems encountered with implementation of data collection and processing procedures, to include missed

opportunities, should be evaluated to determine the causes. The results of the monitoring and evaluation process may well lead to adjustments in procedures. Procedural changes must be documented in the form of updated instructions for data collection and processing, which must then be disseminated to all personnel involved. In addition, it may be advisable to reinforce the priority of the data collection efforts in collaboration with the sponsors in the various domains.

Application Environment

Characteristics of the environments in which personnel operate the database's components and in which customers use the database's products influence the success of the database. It is important to identify obstacles promptly and seek realistic solutions. Even when notable problems are not found, the effectiveness and efficiency of the database can usually be improved by tuning the system to better fit the implementation environment.

Support the Customers

Potential customers will have a vague notion at best regarding how the database can benefit them. Job aids such as information brochures, operational guides, and training packages are important to help customers, especially unit trainers, realize the full potential of the database. These job aids should help the customer understand the capabilities of the database and how the functional features can help them accomplish their jobs. A good example of a promising idea for an operational guide relates to AARs in the SIMNET environment. Without a guide explaining how to use UPAS output to prepare for AARs, unit personnel find it difficult to incorporate the newly available information into their established AAR procedures. Other means--hotlines, electronic bulletin boards, World Wide Web pages, service representatives, training sessions, etc.--can help educate potential customers of the database. Customer feedback can be solicited and used to refine the various means in order to make them more effective.

Support the Data Collectors

Technical support of data collection personnel is a must to maintain the integrity of the data pipelines. Where established data collection procedures are involved, technical support is normally built into the operational training support network. However, when new data collection components are introduced, especially on an experimental basis, special arrangements are most likely required to provide technical support to operators. Troubleshooting procedures and repair guidelines should be documented in writing for the operators. It may be feasible to embed "Help" routines in the workstations for troubleshooting situations. A central support cell can be staffed to respond to queries and help solve problems related to hardware and software. Convenient and rapid channels for accessing the support cell should be defined, driven by the field operating environment. The support cell should be equipped with a duplicate of each component being supported in the field, and should have local access to an operational environment similar to that found at the field sites. Replacement parts and spare machines located on-site can help keep operations on track when failures occur. Planning for logistical support should begin early, normally in parallel with the database design phase.

Maintain Community Support

A valuable mechanism for maintaining interest and support at the various data collection sites is to provide periodic feedback to the supporting agencies. The feedback can summarize data collection accomplishments and milestones to demonstrate the progress being made as well as the value to the participating organizations. Such feedback can serve as an incentive to operators and maintainers, and it can generate interest among potential customers of the database's products.

Support Simulation Terrain Databases

In simulation environments such as SIMNET and CCTT, terrain is emulated in the form of digital terrain databases. The terrain databases must be converted to a special format for use on automated data collection devices such as UPAS. Multiple terrain databases are often used at a given site, and new terrain databases come online occasionally. To support the full spectrum of training exercises, data collection platforms should be equipped with the complete set of terrain databases in use at the site. Initial installation of data collection workstations should include all desired terrain databases. Further, operators should have access to a capability for converting new databases. Placing a conversion package in the hands of operators may be cost effective.

Continuously Improve the Database

The research team should monitor customers' use of database products to identify problems encountered and improvements needed. Further, the team should monitor the reactions and suggestions of personnel operating and maintaining the various database components, to define operational enhancements needed. The goal is to understand the operating environments of customers and operators so their needs can be met and potential obstacles can be overcome. Routine mechanisms for gathering feedback (e.g., customer satisfaction questionnaires, suggestion boxes) are valuable, and they can be supplemented by visits to customer locations and data collection sites. Suggestions for improving and enhancing the various components of the database should be solicited. Program evaluation methods can help gather realistic feedback. An effective monitoring process will lead to definitive steps for modifying the database, with the aim of increasing its value to the combined training community it serves.

Summary of Lessons Learned

Table 4 summarizes the lessons discussed in this section. The summary statements are necessarily simplified for the sake of brevity. For a full discussion of the lessons learned, the reader is referred to the preceding subsections.

Table 4 - Summary of Lessons Learned

Topic	Lesson Title	Lesson Learned
Database Design	Define the Target Audiences	Define clearly the target audience groups: customers, operators, and maintainers.
	Meet Needs of All Target Audiences	Ensure the design process meets needs, preferences, and habits of customers, operators, and maintainers.
	Understand Targeted Data	Obtain descriptions of targeted data, including the data collection environment, as early as possible.
	Ensure Data Collection Complements Training	Design data collection tools so they do not interfere with training.
	Ensure Operator Incentives	Design data collection tools to provide immediate value to operators and their direct customers.
	Plan for Personnel Requirements	Plan early for personnel required to operate and maintain the database components.
	Plan for Logistical Support	Plan early for logistical support of the database components.
	Ensure User Friendly Interfaces	Incorporate graphical user interfaces, unless the targeted operators prefer a text-based interface.
	Plan for Multi-Phase Design Contingency	Plan, schedule, and resource design activities in phases, when necessary.
	Document the Design	Document the database design in a functional description, and update it as changes occur.
Database Development	Obtain Target Audience Input	Obtain input of customers, operators, and maintainers throughout development.
	Sample Targeted Data	Obtain sample data from targeted sources as early as possible.
	Verify Feasibility of MOPs	Routinely crosswalk desired MOPs with viable data elements, and update MOP list as necessary.
	Meet Operator Expectations	Match hardware, software, and operating procedures with expectations and habits of targeted operators.
	Test Components Thoroughly	Test all database components at each stage of development, identifying problems and documenting corrective actions.
	Plan Installation Early	Initiate early coordination with site personnel for installation of new data collection components.
	Have Developers Install Components	Have component developers install new hardware/software whenever possible.

(table continues)

Topic	Lesson Title	Lesson Learned
Data Collection and Processing	Ensure Motivated Data Collectors	Arrange for sponsors to communicate the high priority of collecting and shipping quality data. Educate operators and unit trainers on the training value of the database, including data collection components. Arrange for knowledgeable on-site personnel to routinely supervise data collection activities.
	Standardize the Data	Use standard training packages and performance assessment procedures to standardize data.
	Provide Qualified Operators	Ensure data collection and processing operators are qualified and trained; provide refresher training.
	Protect Quality of Data	Safeguard data quality by routine monitoring of procedures, screening of data, feedback to data collectors, and site visits.
Application Environment	Support the Customers	Educate customers on uses and benefits of the database. Provide technical information and services to help customers utilize the database products.
	Support the Data Collectors	Provide responsive technical and logistical support of data collection operations.
	Maintain Community Support	Provide periodic feedback to supporting/sponsoring agencies and customers re: database accomplishments.
	Support Simulation Terrain Databases	Ensure timely availability of digital terrain databases for simulation-based data collection components.
	Continuously Improve the Database	Maintain proactive database improvement program anchored to feedback from customers, operators, and maintainers.

Recommendations for Future Development

The observations and lessons learned from the current project point to follow-on efforts which have the potential to enhance significantly the Army's training assessment and management capabilities. Building largely on the integrated database capabilities, the following steps are recommended.

1. Upgrade the UPAS workstation as a low-cost tool for the SIMNET environment, emphasizing increased user friendliness and compatibility with the emerging STAARS.
2. Develop UPAS utilization guides for unit trainers, such as a trainer's guide, an AAR support guide, and a unit training package.
3. Expand and accelerate efforts to export SIMUTA exercises for use in SIMNET facilities in Germany.

4. Expand the scope of the integrated database to include data on tactical decision making, command and control, communications, etc., as well as unit training histories, training costs, subjective assessments of training effectiveness, and similar information.

5. Populate the integrated database with routine data from USAREUR: establish and maintain continuing data pipelines, including CMTC data; develop collection-point data screening mechanisms; develop remote data entry capabilities.

6. Create an analysis center to support customer needs: establish database operations and maintenance service, then expand to include analysis capabilities.

7. Demonstrate the potential utility of the integrated database by designing and implementing statistical analyses to answer limited questions of practical value to the Army training community.

8. Educate potential customers regarding the integrated database's capabilities, using such means as World Wide Web pages, videocassettes, and multimedia CD ROMs.

9. Establish a continuing integrated database improvement program, based on feedback from customers, operators, and maintainers.

10. Interface the integrated database with the ATDL, and expand the data capture network to meet the growing needs of the Army training community (e.g., constructive simulation training programs, Leader Training Programs).

11. Establish a library of analytical tools and reports; include periodic reports of general interest, for routine dissemination or online access.

12. Expand data input capabilities to include electronic clipboards, notebook computers, and similar sources; emphasize remote data entry, perhaps via electronic polling.

13. Explore remote access capabilities (e.g., Internet, wide area network) under controlled conditions to empower customers to obtain integrated database information directly.

14. Develop capabilities for interfacing the integrated database with STAARS, SATS, and other training management systems.

15. Explore requirements for expanding the integrated database into the joint training domain.

CONCLUSIONS

A comprehensive unit training performance database is an essential component of the Army's complex training environment. Such a database will provide Army trainers, training managers, training developers, policy makers, and researchers access to valuable information. The integration and sharing of objective performance and resource data can be expected to lead to greater efficiencies in Army training and more effective training outcomes. The prototype integrated database described in this report is a first step toward the establishment of a comprehensive, Army-wide data pool. The recommendations for future research provide a road map for establishing a comprehensive database, including analytical capabilities to support the Army training, research, and development communities.

The lessons learned in this project offer real-world tested guidelines and tips of value to unit training performance database developers. Along with the prototype database, they provide a foundation for expanding the Army's capabilities to collect and manage training performance data. A prominent benefit of an integrated data pool lies in the ability to answer probing questions regarding the cost-effectiveness of new training tools, policies, and techniques. More effective training programs for Army combat units are the ultimate goal of this research, with enhanced mission readiness and cost savings the ultimate reward.

REFERENCES

Alluisi, E.A. (1991). The development of technology for collective training: SIMNET, a case history. Human Factors, 33, 343-362.

Boldovici, J.A., & Bessemer, D.W. (1994). Training research with distributed interactive simulations: Lessons learned from SIMNET (ARI Technical Report 1006). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 285584)

Briscoe, J.A. (1996a). Database functional description (ARI Contractor Report). Manuscript in preparation.

Briscoe, J.A. (1996b). Integrated database system and supporting documentation (ARI Contractor Report). Manuscript in preparation.

Burnside, B.L. (1990). Assessing the capabilities of training simulations: A method and Simulation Networking (SIMNET) application (ARI Research Report 1565). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 226354)

The COBRAS Team. (1995). Force XXI Training Program: Development of virtual and constructive simulation-based training for the conventional mounted brigade--research plan (HumRRO Interim Report IR-PRD-95-13). Alexandria, VA: Human Resources Research Organization.

Drucker, E.H., & Campshure, D.A. (1990). An analysis of tank platoon operations and their simulation on Simulation Networking (SIMNET) (ARI Research Product 90-22). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 226956)

Hancock, W., & Bayha, F. (1982). The learning curve. In Salvendy, G. (Ed.), Handbook of industrial engineering. New York: John Wiley and Sons.

Hiller, J., McFann, H., & Lehowicz, L. (1994). Does OPTEMPO increase unit readiness? In R.F. Holz, J.H. Hiller, & H.H. McFann (Eds.), Determinants of effective unit performance (pp. 71-79). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hoffman, R.G., Graves, C.R., Koger, M.E., Flynn, M.R., & Sever, R.S. (1995). Developing the Reserve Component Virtual Training Program: History and lessons learned (ARI Research Report 1675). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 296153)

Madden, J.L. (1989). Unit training management system, strategy and program for SIMNET (BDM TR BDM/MTY-TR-0028-89). Monterey, CA: BDM International, Inc.

Meliza, L.L., Bessemer, D.W., Burnside, B.L., & Shlechter, T.M. (1992). Platoon-level after action review aids in the SIMNET unit performance assessment system (ARI Technical Report 956). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 254904)

Meliza, L.L., Tan, S.C., White, S., Gross, W., & McMeel, K. (1992). SIMNET unit performance assessment system (UPAS) user's guide (ARI Research Product 92-02). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA 250699)

Meliza, L.L., Tan, S.C., White, S., Gross, W., & McMeel, K. (1994). SIMNET unit performance assessment system advanced user's guide (ARI Research Product). Manuscript in preparation.

U.S. Army Training and Doctrine Command. (1994). Force XXI operations: A concept for the evolution of full-dimensional operations for the strategic Army of the early twenty-first century (TRADOC Pamphlet 525-5). Fort Monroe, VA: Author.

U.S. Department of the Army. (1989). Standard Army Training System (SATS) menu/screen dictionary. Fort Monroe, VA: Headquarters, U.S. Army Training and Doctrine Command.

Wunsch, B.J., & Cassidy, J.J. (1996). Unit performance assessment system (UPAS) operating guide (ARI Contractor Report). Manuscript in preparation.

ABBREVIATIONS

AAR	After Action Review
ARI	U.S. Army Research Institute for the Behavioral and Social Sciences
ATAFS	Automated Training Assessment and Feedback System
ATDL	Army Training Digital Library
CALL	Center for Army Lessons Learned
CCTT	Close Combat Tactical Trainer
CD ROM	Compact Disc Read Only Memory
CMTC	Combat Maneuver Training Center
CMTC-IS	Combat Maneuver Training Center-Instrumentation System
COBRAS	Combined Arms Operations at Brigade Level, Realistically Achieved through Simulation
CONUS	Continental United States
CPX	Command Post Exercise
CTC	Combat Training Center
DIS	Distributed Interactive Simulation
DMDC	Defense Manpower Data Center
DoD	Department of Defense
ECI	Electronic Collection Instrument
FTX	Field Training Exercise
ITMS	Integrated Training Management Strategy
JRTC	Joint Readiness Training Center
MOP	Measure of Performance
NTC	National Training Center
OPORD	Operation Order
PC	Personal Computer
SATS	Standard Army Training System
SIMNET	Simulation Networking
SIMUTA	Simulation-Based Multiechelon Training for Armor Units
STAARS	Standard Army After Action Review System
TADSS	Training Aids, Devices, Simulators, and Simulations
TSP	Training Support Package
UCOFT	Unit Conduct of Fire Trainer
UPAS	Unit Performance Assessment System
USAREUR	U.S. Army, Europe
7ATC	7th Army Training Command

APPENDIX A

MEASURES OF PERFORMANCE

Measure of Performance	Source
<p>----- HOME STATION TRAINING -----</p>	
Tank Unit Conduct of Fire Trainer (UCOFT) Matrix level completed	7ATC Files
Bradley Unit Conduct of Fire Trainer (UCOFT) Matrix level completed	7ATC Files
TOW/DRAGON Gunnery Training Target range	7ATC Files
Firing outcome (hit/miss/misfire)	7ATC Files
<p>----- LOCAL/MAJOR TRAINING AREAS -----</p>	
Platoon Gunnery Trainer (Tank) Avg opening time, vehicle targets	7ATC Database
Avg opening time, troop targets	7ATC Database
Avg kill time, vehicle targets	7ATC Database
Avg kill time, troop targets	7ATC Database
Percent Red vehicle targets killed	7ATC Database
Percent troop targets killed	7ATC Database
Number rehits after kills, all targets	7ATC Database
Percent Blue targets killed (frat)	7ATC Database
Score: Percent total targets killed	7ATC Database
Qualification status	7ATC Database
Platoon Gunnery Trainer (Bradley) Avg opening time, vehicle targets	7ATC Database
Avg opening time, troop targets	7ATC Database
Avg kill time, vehicle targets	7ATC Database
Avg kill time, troop targets	7ATC Database
Percent Red vehicle targets killed	7ATC Database
Percent troop targets killed	7ATC Database
Number rehits after kills, all targets	7ATC Database
Percent Blue targets killed (fratricide)	7ATC Database
Score: Percent total targets killed	7ATC Database
Qualification status	7ATC Database

Measure of Performance**Source**

Tank Gunnery Table VIII

Number of main gun first round hits	7ATC Database
Total number of main gun rounds fired	7ATC Database
Total number of main gun targets hit	7ATC Database
Total number of troop targets hit	7ATC Database
Total score (sum of 10 tasks)	7ATC Database
Total main gun opening time	7ATC Database
Total main gun closing time	7ATC Database
Machine gun opening time	7ATC Database
Total machine gun closing time	7ATC Database

Tank Gunnery Table XII (per crew)

Percent main gun targets hit	7ATC Database
Percent troop targets killed	7ATC Database
Percent total targets hit	7ATC Database
Main gun score	7ATC Database
Ratio of main gun hits/firings	7ATC Database
Total score	7ATC Database
Qualification status (platoon)	7ATC Database

Bradley Gunnery Table VIII

Total engagement time, main gun	7ATC Database
Total engagement time, coax gun	7ATC Database
Kill time, main gun	7ATC Database
Kill time, coax gun	7ATC Database
Total number of targets hit, main gun	7ATC Database
Total number of targets hit, coax gun	7ATC Database
Total crew score	7ATC Database

Bradley Gunnery Table XII (per crew)

Percent main gun targets hit	7ATC Database
Percent troop targets hit	7ATC Database
Percent total targets hit per band	7ATC Database
Main gun score	7ATC Database
Troop target score	7ATC Database
Total score	7ATC Database
Qualification status (platoon)	7ATC Database
Percent main gun targets hit (Offense)	7ATC Database
Main gun score (Offense)	7ATC Database
Percent main gun targets hit (Defense)	7ATC Database
Main gun score (Defense)	7ATC Database

Measure of Performance	Source
Aviation Crew Gunnery Tables VII and VIII	
Opening time	7ATC Database
Total task time	7ATC Database
Number of bursts fired	7ATC Database
Number of rounds in box	7ATC Database
Number of target hits	7ATC Database
Crew score	7ATC Database
Crew error penalty points	7ATC Database
Task outcome (GO/NO GO)	7ATC Database

----- SIMULATION NETWORKING (SIMNET) -----

Direct Fire Engagement	
Number of Blue rounds fired, by vehicle type and ammo type	UPAS
Number of Blue hits, by vehicle type and ammo type	UPAS
Number of Blue kills, by vehicle type and ammo type	UPAS
Total number of Blue rounds fired, by vehicle type	UPAS
Total number of Blue hits, by vehicle type	UPAS
Total number of Blue kills, by vehicle type	UPAS
Firing range and result, by individual round (BLUFOR)	UPAS
Number of Blue rounds fired, by individual vehicle	UPAS
Number of Blue impacts, by individual vehicle	UPAS
Blue firing activity, by individual vehicle: ID, range band, # rounds fired, percent hits, minimum range, maximum range	UPAS
Number of Red rounds fired, by vehicle type and ammo type	UPAS
Number of Red hits, by vehicle type and ammo type	UPAS
Number of Red kills, by vehicle type and ammo type	UPAS
Total number of Red rounds fired, by vehicle type	UPAS
Total number of Red hits, by vehicle type	UPAS
Total number of Red kills, by vehicle type	UPAS
Blue fratricide events: ID, time, target, result, range	UPAS
Fire Support	
Indirect fire damage: ID, time, side, outcome	UPAS

Measure of Performance	Source
Miscellaneous	
Number of Blue vehicles alive at start of exercise	UPAS
Number of Blue vehicles alive at end of exercise	UPAS
Number of Blue vehicles dead at end of exercise	UPAS
Number of Red vehicles alive at start of exercise	UPAS
Number of Red vehicles alive at end of exercise	UPAS
Number of Red vehicles dead at end of exercise	UPAS
Blue crew error events: ID, time, outcome (e.g., collision)	UPAS
----- COMBAT MANEUVER TRAINING CENTER (CMTC) -----	
Direct Fire Engagement	
Number of BLUFOR rounds fired, by wpn system	CMTC-IS
Number of BLUFOR hits, by wpn system	CMTC-IS
Number of BLUFOR kills, by wpn system	CMTC-IS
Percent BLUFOR hits, by wpn system	CMTC-IS
Percent BLUFOR kills, by wpn system	CMTC-IS
Number of BLUFOR rounds fired per kill	CMTC-IS
Number of BLUFOR vehicle losses, by system	CMTC-IS
Engagement range (interval), by target type	CMTC-IS
Engagement range (interval), by wpn system	CMTC-IS
Engagement range (interval), by unit	CMTC-IS
Number of Blue personnel casualties, by unit	CMTC-IS
Number of Blue personnel casualties, by category	CMTC-IS
Fire Support	
Number of Blue fire mission requests	CMTC-IS
Number of Blue fire mission requests not executed	CMTC-IS
Number of Red losses due to arty, by veh type	CMTC-IS
Number of Red losses due to mortars, by veh type	CMTC-IS
Number of Blue arty fratricides, by veh type	CMTC-IS
Number of Blue mortar fratricides, by veh type	CMTC-IS
Number of Blue arty rounds fired, by type	CMTC-IS
Number of Blue mortar rounds fired, by type	CMTC-IS
Battle Command	
Number of radio transmissions	CMTC-IS
Avg length of radio transmissions	CMTC-IS
Number of radio transmissions 25-55 sec long	CMTC-IS
Number of radio transmissions >55 sec long	CMTC-IS

Measure of Performance	Source
Combat Service Support	
Percent operational readiness	CMTC-IS
Number of replacement personnel requisitioned	CMTC-IS
Number of rounds requisitioned, by type	CMTC-IS
Number of replacement vehicles requisitioned	CMTC-IS
Event Logs	
Fratricide log	CMTC-IS
Obstacle effectiveness	CMTC-IS
Elements of information, by category	CMTC-IS
Elements of information, by unit	CMTC-IS
BLUFOR fire support log	CMTC-IS
BLUFOR fire support results	CMTC-IS
Chemical casualty log	CMTC-IS